Bio-inspired Based Techniques for Thermogram Breast Cancer Classification

Ammar Abdulrahman Ahmed[1]*   **Mona A. S. Ali[1,2]**   **Mazen Selim[1]**

[1]Department of Computer Science, Faculty of Computers and Informatics,
Benha University, Qalubia, Egypt
[2]Department of Computer Science, College of Computer Science and Information Technology,
King Faisal University, Kingdom of Saudi Arabia
* Corresponding author's Email: am30.ah30@gmail.com

Abstract: Nowadays, breast cancer considered a main cause of death for women all over the world. It is defined as a group of cells that grow rapidly and causes the formation of a lump in breast tissue which leads to tumor formation which can be categorized either malignant (cancerous) or benign (non-cancerous). On the other side, mammography as a screening and diagnostic tool suffers from some limitations, especially with young women who have dense breasts. Therefore, there was a need to develop more effective tools. Thermography is an imaging tool used to record the thermal pattern. The main contribution of this paper is proposing a unique method for classifying the breast thermography images into one of three classes: normal, benign, or malignant. Additionally, bio-inspired algorithms namely, ant colony optimization (ACO) and particle swarm optimization (PSO) are used for feature selection. The proposed method contains four phases: Image preprocessing, feature extraction, feature selection, and classification. The proposed method is assessed using a benchmark thermography dataset. The experimental results show that our method has a promising performance.

Keywords: Machine learning, Breast cancer classification, Thermography, ACO, PSO, SVM.

1. Introduction

Breast cancer can be considered the most common kind of cancer among women [1]. It is defined as a set of cells of fast growth causing an extra mass formation in breast tissue which leading to the formation of tumors. Generally, Those Tumors are being classified into malignant (cancerous) or benign (non-cancerous) [2]. Benign tumors are not dangerous on the person's life; however, it can increase the chance of getting a breast cancer. On the other side, malignant tumors are dangerous on the person's life because of their cancerous nature. However, many studies have concluded that detecting malignant tumors in an early stage can significantly increase the survival rates [3].

Survival rates because of breast cancer differ clearly among the countries of the world. However, it noticed that the survival rates are particularly low in the developing and low-income countries due to the absence of early detection programmes of breast cancer. Unfortunately, physicians consider breast cancer a heterogeneous disease which means that the early detection process of this disease is considered a major clinical challenge [1].

Currently, breast tumors can be detected and diagnosed by employing different imaging tools such as thermography and mammography. Thermography technique uses a camera with thermal infrared sensitivity to capture the distribution of temperature in the human body. Unlike mammography and other imaging modalities, thermography is a passive and harmless technique without radiations. Likewise, it isn't invasive, does not need direct contact with the human body, quick, and need minimal effort. Thermal imaging has been shown to be a suitable option for detecting the blood perfusion changes which may happen because of angiogenesis, inflammation, or other reason. Therefore, thermography is able to detect the

existence of tumors especially when the tumor is in dense tissue or in its early stage [4]. In addition, unlike mammography which able to detect the tumors only when its size exceeds a certain limit, thermography can even detect the small tumors where the regions that contain a tumor are usually having higher temperature when compared with the normal regions [5]. Consequently, thermography is expected to be one of the most effective tools for carcinoma detection in an early stage.

Nowadays, the field of automated diagnosis systems development is getting an increasing attention from the researchers to help the physicians in the decision making process. On the other hand, machine learning methods are playing a vital role in building such systems, where they can be employed to perform many tasks such as feature selection and classification [2]. Although many research works have been proposed in order to characterize the breast masses, there is a big room for improving the accuracy. In addition, these works only classify breast masses into normal/abnormal or benign/malignant. In this paper, a breast cancer classification method, which depends on thermal breast images, is suggested. The main steps of the proposed system are listed below.

1) *Image preprocessing* which includes Region-of-Interest (ROI) segmentation, resizing, RGB to grey scale conversion, and applying the median filter for noise removal.
2) *Feature extraction* in which two groups of features are extracted including Grey Level Co-occurrence Matrix (GLCM) based features and Bag-of-Features.
3) *Feature selection* in which the most discriminant features are selected to accelerate the classification process. In the proposed work, we use two well-known bio-inspired techniques namely, ant colony optimization (ACO) and particle swarm optimization (PSO) for feature selection.
4) *Classification* in which a Support Vector Machine (SVM) based classifier is used to classify the tumors into one of three classes: normal, benign, or malignant.

The proposed system uses a hybrid feature vectors that includes both of statistical features and texture features. Also, it employs bio-inspired approaches to perform the feature selection step in order to reduce the training time, increase the classification accuracy, and overcome the over-fitting problem. In addition, it turns the problem into a multi-class problem by classifying the breast masses into normal, benign, or malignant. The proposed system is evaluated using a standard thermal breast images dataset. The obtained results show that the proposed system has a better performance in terms of a number of performance metrics compared to other recently proposed systems.

The remaining sections of the paper are organized as follow: Section 2 presents some of the previous research efforts which have been performed in the field. Section 3 introduces the proposed method. Section 4 presents the dataset description and the experimental results. Finally, Section 5 contains the conclusion and future work.

2. Related work

Due to the importance of breast cancer, as one of the most common diseases for women, a lot of research efforts have been done in the last years to detect this disease in an early stage. Recently, mammography and thermography are considered the most used modalities in breast cancer detection [6]; therefore, this section starts with reviewing some of the recent works that employ mammogram images to characterize the breast masses [7 - 10].

Rouhi *et al.* have proposed two methods to classify breast masses in mammograms as benign or malignant. Both methods involve pre-processing, Region of Interest (ROI) segmentation, feature extraction, feature selection, and classification. To perform the segmentation process, the first method uses an automated region growing while the second method uses a Cellular Neural Network (CNN). Feature extraction step in both methods extract various features including intensity, textural, and shape features from the segmented images. Also, the feature selection step is performed in both methods using the Genetic Algorithm (GA). Different classifiers are employed to perform the classification process including random forest, naïve Bayes, Support Vector Machine (SVM), and k-Nearest Neighbour (KNN). Two datasets are used to evaluate the performance of the proposed methods namely, mammography imaging analysis society (MIAS) and digital database for screening mammography (DDSM). A number of experiments are conducted and the best results obtained are 96.87% sensitivity, 95.94% specificity, 96.47% accuracy [7].

Vaidehi *et al.* have presented an automated system for helping the radiologist to analyze breast masses in mammogram images. The proposed system uses the fuzzy C-means technique to perform the segmentation process. Then, Grey Level Co-

occurrence Matrix (GLCM) texture features are extracted in the feature extraction step. Finally, three classifiers namely, adaboost, back propagation neural network and sparse representation classifiers are used to classify breast masses as benign or malignant. The results obtained using mammographic image analysis society (MIAS) database images show the superiority of the sparce representation classifier with 90.62% sensitivity, 99.99% specificity, and 93.75% accuracy [8].

Taheri *et al.* have proposed a method to distinguish between normal and abnormal breast mammogram images. The proposed method involves image pre-processing step in which the noise and unwanted objects are removed. Then, the intensity value, auto-correlation matrix value of detected corners, and energy are extracted to train a Support Vector Machine (SVM) classifier. The proposed method is evaluated using a dataset that consists of 600 mammogram images (300 benign and 300 malignant). The obtained results show that the proposed method has 96.8% precision and 92.5% sensitivity [9].

Further work is presented by **Thawkar et al.** to automatically detect and classify breast masses as benign or malignant. The proposed method uses a number of techniques to automatically detect the suspicious masses including Otsu's global thresholding technique, morphological operations and watershed transformation. In the feature extraction step, twenty-five features based on intensity, texture and shape are extracted. Then, a feature selection step is performed using the step-wise Linear Discriminate Analysis in order to determine the eight most discriminating features. Finally, a number of classifiers are employed to perform the classification process including Fisher's Linear Discriminate Analysis, Support Vector Machine and Multilayer Perceptron (MLP) with two training algorithms Levenberg-Marquardt and Bayesian Regularization. The results obtained on 651 mammogram images show the superiority of MLP-RBF over other classification techniques with 98.9% accuracy and area under Receiver Operating Characteristics curve AZ=0.98±0.007 [10].

In spite of the good results obtained in [7 - 10], there are two main disadvantages relevant to these works. First, they use the mammograms in order to characterize breast masses and mammography has many disadvantages such as being invasive and the electromagnetic radiation itself leads to the growth of breast tumor by 2 % during every periodic inspection [1]. Second, the previous works aimed for only classifying breast masses into normal/abnormal or benign/malignant. It would be better if a system is presented that is able to classify breast masses into normal/abnormal and further classify abnormal masses into benign/malignant using thermograms. Many works have depended on thermograms in order to overcome the problems of mammograms [1, 4, 5, 11 - 16].

Sayed *et al.* [1] aimed at measuring the effectiveness of four bio inspired swarm algorithms including Particle Swarm Optimization (PSO), Grey Wolf Optimization (GWO), Moth Flame Optimization (MFO), and Firefly Algorithm (FA) to handle the biomedical images segmentation issue and measuring the ability of the selected swarm intelligence algorithms in the classification phase. The proposed system extracted different types of features including statistical and texture features. Then, a Support Vector Machine classifier is used to determine if a given thermogram contains a benign or malignant breast mass. The obtained results have revealed that FA algorithm provides the best results with 96.83 % accuracy, 94.87 % sensitivity, 100 % specificity and 100 % precision.

Another approach for classifying thermal breast images is presented by **Krawczyk et al.** [4]. The proposed approach uses statistical image features and an efficient ensemble method for imbalanced classification problems. In the proposed approach, a set of features are obtained from the thermography images to detect asymmetry between left and right breast regions. Then, the obtained features are used as input to a hybrid classifier that is built based on support vector machines and neural networks. The experiments have been done on a breast thermography images, experimental results have revealed that the proposed approach is doing well with imbalanced data with 89.03% accuracy, 81.96% sensitivity, and 90.8% specificity in their best case. Further work presented by **Krawczyk et al.** [11] in which they extract features that can describe bilateral symmetries in thermography images and then utilize an ensemble classifier to decide. Their classification technique addressed the issue of imbalanced class distribution through constructing feature subspaces from balanced data subsets and training the different classifiers on different subspaces. The fusion process among the different classifiers is done in two ways. In the first way, each classifier takes a weight that is dynamically assigned using an evolutionary algorithm, while the second way performs the fusion process using a neural network. The best results have been achieved by the fusion performed using the neural network with 90.03% accuracy, 80.35% sensitivity, and 90.15% specificity.

Another relevant work that adopts a number of machine learning algorithms for breast cancer early detection based on thermal breast images is provided by **Oleszkiewicz et al.** [12]. The classification process is done using a Support vector machine (SVM) classifier that has been trained using SVM and Sequential Minimal Optimization algorithm (SMO) techniques. Also, a heuristic method for tuning the parameters of the SVM classifier was presented. The proposed method has been evaluated using the sensitivity and specificity performance metrics on a breast thermography images dataset.

Schaefer et al. [13] used a set of statistical features derived from the comparison between left and right breast regions. Then, these features are passed to a diagnosis system that depends on fuzzy rules. The used classifier extracts fuzzy if-then rules using a training set that comprises known cases. A number of experiments have been conducted using a dataset that consists of nearly 150 cases and the obtained results have revealed that the proposed system has accuracy of 80%.

Also, **Pramanik et al.** [14] have presented a local feature extraction module called block variance (BV) to separate among malignant and benign thermal breast images. The proposed method determines the contrast texture in the grey-scale thermal images based on the local variation of intensities. Also, the proposed method identifies abnormality using asymmetric temperature distribution between right and left breast images. Also, they have used a feed-forward neural network for classification. The proposed system has been assessed using thermal images obtained from a database called "Database for Mastology Research (DMR)" and has achieved classification accuracy of 90%.

In addition, **Gaber et al.** [5] have developed a Computer-Aided Detection (CAD) system that can differentiate between normal and abnormal thermal breast images. The proposed approach includes two stages: segmentation and classification. In the first step, a segmentation method that depends on optimized Fast Fuzzy c-mean (F-FCM) algorithm and Neutrosophic sets (NS) has been suggested. In the second step, a SVM classifier has been trained using different kernel functions. A number of experiments have been conducted using a benchmark database for evaluating the performance of their approach. The obtained results have revealed that the proposed system has an accuracy of 100%.

Sathish et al. [15] have proposed an automated method for segmenting right and left breast to analyze breast thermogram images for breast cancer detection. The proposed technique depends on the breast's shape features as well as the polynomial curve fitting. The proposed system extracts histogram and GLCM features from the segmented image. Then, it employs a SVM classifier with Radial Basis Function (RBF) kernel function to differentiate between normal and abnormal thermal breast images. The proposed method has been assessed using a public breast thermogram images database that consists of fourty normal and fourty abnormal thermogram images. The obtained experimental results have shown an accuracy of 90%.

Finally, **Hossam et al.** [16] have introduced another automated segmentation method to determine the Region of Interest (ROI) image in breast thermograms using the images statistics. Their system starts with detecting the boundaries of ROI using the proposed method. Then, the detected boundaries are enhanced using the Hough transform algorithm and another approach proposed for enhancing the contrast of the detected boundaries. The proposed system extracts statistical features from the ROI images and employs two classifiers to differentiate between normal and abnormal images, namely SVM classifier and Artificial Neural Network (ANN) classifier. The proposed system has been evaluated using breast thermograms database called DMR-IR and has achieved 96.67% and 96.07% for the SVM classifier and the ANN classifier, respectively.

Despite of the good results obtained in the previous works that use thermograms to characterize the breast masses. All of these works classify breast masses into normal/abnormal or benign/malignant. Therefore, the objective of the paper is to propose a system that is able to classify breast masses into normal/abnormal and further classify the abnormal breast masses into benign or malignant masses.

3. The proposed work

In this section, a thermography-based breast cancer classification method is proposed to classify the thermography images into one of three categories: normal, malignant or benign. As depicted in Fig. 1, the proposed system consists of four steps: image preprocessing, feature extraction, feature selection, and classification. A detailed description for each step is given in the following subsections.
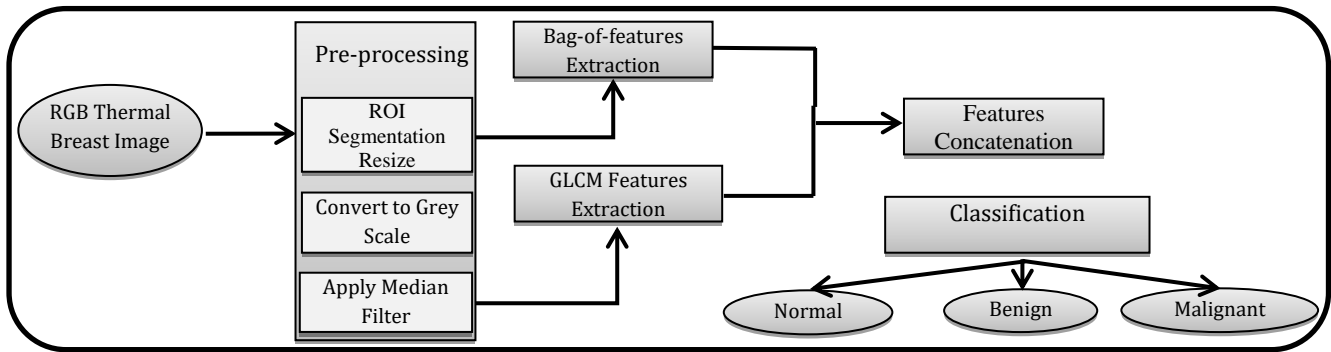
Figure. 1 Pictorial diagram of the proposed method using feature selection

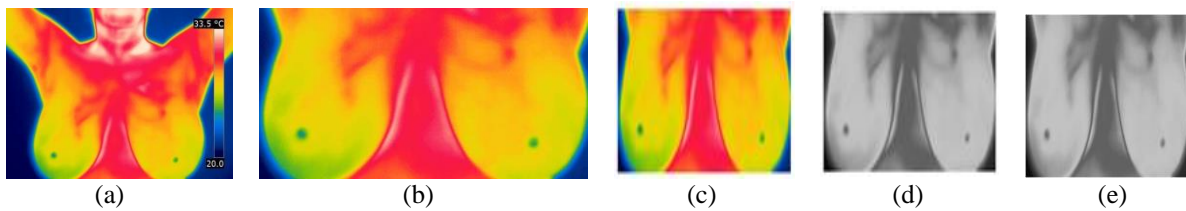| (a) | (b) | (c) | (d) | (e) |

Figure. 2 The different stages of preprocessing step: (a) original image, (b) segmented image, (c) resized image, (d) grey scale image, and (e) after apply median filtered image

3.1 Image preprocessing

In the preprocessing step, a set of operations are applied on the thermal breast images including: Region of Interest (ROI) segmentation, resizing, converting the thermal images from RGB into grey scale, and applying the median filtering. The used thermal breast images include the breast region in addition to some extra body parts, as shown in Fig. 2 (a). A prepressing step is applied to make them suitable for the feature extraction step.

In this work, with the help of a medical expert, the region that contains the right and left breasts is segmented manually from the colored thermal images as shown in Fig. 2 image (b). In order to unify the size of the segmented ROIs, a resizing process is applied to make the size of all segmented ROIs 128x128, as shown in Fig. 2 image (c). The resized image is changed over into grey scale image as shown in Fig. 2 image (d). Converting the segmented ROIs into equivalent grey scale images is necessary for the feature extraction step (GLCM features).

Finally, the last operation during this stage is applying the median filter using a window of 3 x 3 in order to remove or reduce the noise in the images while preserving the edges to enhance the results of the subsequent steps, as shown in Fig. 2 image (e).

3.2 Feature extraction

Two feature extraction approaches are employed to extract a group of discriminating features that can

be effectively used by the proposed system to distinguish among normal, benign, and malignant breast thermal images. The used feature extraction approaches include Bag-of-features and GLCM based features.

3.2.1 Bag of features

Bag of features is a feature extraction approach that can be used for extracting discriminating features that can be used as input for the image classification process. It is also known as bag-of-words techniques. It has been obtained from the natural languages processing filed to be used in the computer vision field. Given a set of descriptors that represent a number of images, a subset of N of them are selected and referred to as a "dictionary". At that point all descriptors can be mapped to the nearest word in that dictionary, allowing these images to be described using the n dimensional vector of word frequencies [17].

In the proposed work, bag of words approaches has been used to extract discriminating features that can be utilized to recognize the different thermal breast images (normal, benign, and malignant). The bag of words approach consists of two main steps: extracting the key points and feature descriptors from the images and quantizing the key points and descriptors using K-means clustering approach to build the visual dictionary [18].

In order to extract the keypoints and calculating the descriptors of these points, two main algorithms

Table 1. The extracted GLCM based features

GLCM Features		
(1) Energy	(9) Cluster Shade	(17) Difference entropy
(2) Entropy	(10) Cluster Prominence	(18) Information measures of correlation 1
(3) Dissimilarity	(11) Maximum probability	(19) Information measures of correlation 2
(4) Contrast	(12) Sum of Square	(20) Maximal correlation coefficient
(5) Inverse difference	(13) Sum Average	(21) Inverse difference normalized (IDN)
(6) Correlation	(14) Sum Variance	(22) Inverse difference moment normalized (IDMN)
(7) Homogeneity	(15) Sum Entropy	
(8) Autocorrelation	(16) Difference variance	

have been employed in the literature, namely scale invariant feature transform (SIFT) and speeded-up robust features (SURF). The major difference between SIFT and SURF the way adopted to implement the space-scale. In SIFT, the space-scale is implemented through iteratively convolving the input image with Gaussian kernel and repeatedly sub-sampled (reduced in size). On the hand, SURF implements the scale-space through applying kernels of increasing sizes to the original image. Another difference is that SIFT descriptor is of 128 dimensions while SURF descriptor is of 64 dimensions [18].

In the proposed work, we have employed the SURF detector to find the key points and to describe the area around these points as a feature vector. SURF approach can be used to extract scale and rotation invariant features in which a set of key points are extracted and a set of 64-dimensional descriptors are constructed for every key point. To extract the keypoints, SURF detector depends on an exceptionally essential exact approximation of Hessian determinant utilizing the Difference-of-Gaussian (DoG) [18].

During the construction of the visual dictionary using the K-means clustering approach, the number of clusters (k) is set to 500. Finally, the feature vectors of the thermal breast images are mapped to the 500 visual words. In the proposed work, the size of each feature vector for each thermal breast image is 500.

3.2.2 Grey level co-occurrence matrix (GLCM) features

In this work, 22 texture features are calculated using Haralick GLCM with the distance parameter d = 1 [19]. The GLCM based computed features are shown in Table 1.

After extracting the different types of features, they are serially concatenated in a larger feature vector of size 522 (500 bag of features and 22 GLCM based features.)

3.3 Feature selection using bio-inspired algorithms

Feature vectors of high dimensionality have many problems such as reducing the classification accuracy and increasing the computational cost. Therefore, the feature selection step is an important step to reduce the dimensionality of the feature vectors which may lead to increasing the accuracy, decreasing the time needed for the training process, decreasing the computational cost, and avoiding the over-fitting problem [11, 20].

Feature selection converts the original feature space into a smaller subspace by selecting the most relevant features. Many feature selection algorithms have been used to implement this task. In this paper, two well-known swarm optimization algorithms, the Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) are used for feature selection. Bio-inspired swarm optimization algorithms are a class of algorithms that are inspired from the collective behavior of animals, bugs and birds [1].

3.3.1 Feature selection using ant colony optimization (ACO)

Ant colony optimization (ACO) is one of swarm intelligence algorithms that have been employed successfully in many optimization problems. It was proposed by M. Dorigo in 1990s. It is inspired from the foraging behavior of ant colonies where ants find the shortest path to food source with the assistance of smelly trails called pheromone [21].

In this paper, the ACO algorithm is employed to find the best subset of features in the dimensional feature vector (522 features) in order to improve the classification accuracy using a number of artificial ants. The ACO algorithm is applied iteratively where three main steps are executed in each iteration [22]:

1. Creating Artificial ants for each subset of features where feature are assigned to artificial

ants using the transition probability function shown below.

$$p_i(t) = \frac{(\tau_i(t))^\alpha {\eta_i}^\beta}{\sum_i (\tau_i(t))^\alpha {\eta_i}^\beta} \qquad (1)$$

where $\tau_i(t)$ is the pheromone value associated with feature i in time t, ${\eta_i}^\beta$ is a heuristic information that reflects the desirability of picking feature i, and α and β are the relative weights of pheromone value and heuristic information, respectively. In the first iteration, all pheromone values and heuristic information values are set to 1 where all the features have equal probabilities to be selected.

2. Evaluating the performance of each ant via measuring the classification accuracy using SVM classifier.
3. Updating pheromone values based on evaporation constant and classification accuracy where the pheromone value τ_i associated feature i is updated using the equation defined below.

$$\tau_i(t + 1) = \rho\tau_i(t) + \Delta\tau_i(t) \qquad (2)$$

where ρ is the evaporation constants whose value is bounded by [0, 1] and $\Delta\tau_i$ indicates the intensification rate which depends on the classification accuracy as shown below.

$$\Delta\tau_i = \frac{1}{(1 - classification\ Accuracy)} \qquad (3)$$

where the classification accuracy is the number of correctly classified patterns divided by the total number of patterns.

3.3.2 Particle swarm optimization (PSO)

Particle swarm optimization (PSO) is a swarm optimization algorithm proposed by Eberhart and Kennedy in 1995[20]. It impressed by the social forage behavior of some animals like flocking behavior of birds and also the schooling behavior of fish. It employs a number of particles that simulate a swarm population moving around in the search space to find the best solution. Each particle provides elect solution for the problem and is usually represented as a point in D-dimensional space. The position of each particle is represented as a vector $x_i = (x_{i1}, x_{i2}, ..., x_{iD})$. Particles moves in the search space to find optimal solutions and each particle has a velocity represented as a vector $v_i = (v_{i1}, v_{i2}, ..., v_{iD})$.

During movement, the position of each particle is updated based on the best position that has been achieved by the particle itself (*pbest*) and the best position that has been achieved by all the particles so far (*gbest*). The velocity and the position of each particle are updated as shown below [23].

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \qquad (4)$$

$$v_{id}^{t+1} = w \times v_{id}^t + c_1 \times r_1 \times (p_{id} - x_{id}^t) \\ + c_2 \times r_2 \times (p_{gd} - x_{id}^t) \qquad (5)$$

where t refers to the iteration number, w is inertia weight that aims at determining the effect of previous velocities on current velocity, c_1 and c_2 are acceleration constants, r_1 and r_2 are random variables whose values are normally distributed in [0,1], and p_{id} and p_{gd} indicate the elements of *pbest* and *gbest* in the dth dimension, respectively

In this paper, the PSO algorithm is employed to find the best feature subset that maximizes the classification accuracy. Toward achieving this task, the binary PSO algorithm has been used [24].

4. Classification

The purpose of this step is to build a classifier that can distinguish between the different classes (normal, benign, and malignant) accurately. In order to build the classifier, we have used a well-known machine learning called support vector machine (SVM) which is a supervised learning algorithm appropriate for data classification and regression. It had been planned by Boser et al. and Vapnik. SVM algorithm is utilized to discover a hyperplane that isolates the classes minimizing training error and maximizing the margin in order to increase generation capability [3]. However, the standard SVM works on two classes only while the proposed work needs multi-class classifier. So, we have used a modified version of the standard SVM called Multi-Class Support Vector Machine [25] for building a classifier that is able to distinguish between many classes. SVM uses vectors x to classify according to the Eq. (6).

$$c = \sum a_i\, k\,(s_i, x) + b \qquad (6)$$

where s_i are the support vectors, a_i are the weights, b is the bias, and k is the kernel function. The constructed classifier uses the polynomial kernel function that can be computed using Eq. (7).
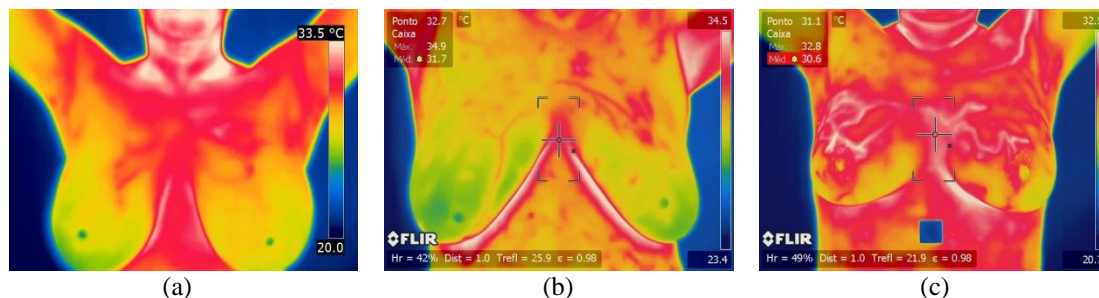
Figure. 3 A sample of the used thermal breast images dataset: (a) normal, (b) benign, and (c) malignant

$$K(x_i, x_j) = (x_i . x_j + 1)^d \qquad (7)$$

where d is the degree of the polynomial. If $c \geq 0$, then x is classified as a member of the primary cluster, otherwise it's classified as a member of the second cluster.

5. Experiments and results analysis

In this section, a set of experiments are performed to evaluate the performance of the proposed method using a thermography breast images dataset in terms of common performance metrics including the classification accuracy, sensitivity, specificity, and recognition time. The equations of the used performance metrics are given below.

$$Accuracy\ (i) = \frac{TP(i)+TN(i)}{TP(i)+TN(i)+FP(i)+FN(i)} \times 100\% \qquad (8)$$

$$Sensitivity(i) = \frac{TP(i)}{TP(i)+FN(i)} \times 100\% \qquad (9)$$

$$Specificity(i) = \frac{TN(i)}{TN(i)+FP(i)} \times 100\% \qquad (10)$$

where $TP(i)$ is the true positive of the ith class, $TN(i)$ is the true negative of the ith class, $FP(i)$ is the false positive of the ith class, and $FN(i)$ is the false negative of the ith class.

5.1 Dataset description

In the proposed work, our experiments have been performed on 118 frontal view breast thermal images of DMR (Database for Mastology Research) database [26]. The DMR consists of normal, benign, and malignant breast thermal images of 287 patients and made public under the approval of the ethical committee where every patient should sign consent.

However, the proposed system has been validated using 118 IR single breast images of size 640 x 480 (30 normal, 45 benign, and 43 malignant). A sample of the different sorts of the used images is shown in Fig. 3.

5.2 Experiments

In this section, a set of experiments have been done to evaluate the performance of the proposed method in terms of classification accuracy, sensitivity, specificity, and recognition time. During the experiments, 70% of the dataset is used for training and 30% of the dataset is used for testing to measure the performance of the proposed method.

In order to setup the associated percentage of the selected features, many values have been adopted starting with step value 10% and ending with 90%. The results have been improved with increasing the selection percentage until 30% for PSO and 40% for ACO and then the obtained results have been deteriorated with increasing the selection percentage. Therefore, the optimal tradeoff between the accuracy and number of features is achieved at 30% of features for PSO and 40% of features for ACO. The recorded values have been chosen according to the optimal selected features from the original feature vector size.

All experiments are done using the same machine with CPU Intel(R) core (TM) i7-3720QM CPU@2.60GHz, 16 GB installed memory (RAM), and 64-bit Windows 10 Pro operating system. The proposed methods are implemented using Matlab 2017a.

First, the proposed system (original system) is evaluated regarding the used performance metrics without applying the feature selection step. Then, the proposed system is evaluated twice, once with applying the feature selection step using ACO referred to as ACO-MSVM and once with applying the feature selection step using PSO referred to as PSO-MSVM.

Table 2. A comparison of the proposed work and other recently proposed methods

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	Recognition Time (Sec.)
Original System	97.14	98	98.6	0.02343
ACO-MSVM	94.29	94.3	97.3	0.00685
PSO-MSVM	97.14	98	98.6	0.00601
Pramanik *et al.* [14].	90	95	85	N/A
Sathish *et al.* [15]	90	87.5	92.5	N/A
Hossam *et al.* [16]	96.667	N/A	N/A	N/A

Also, the proposed work is compared to the works proposed by **Pramanik *et al.*** [14], **Sathish *et al.*** [15], and **Hossam *et al.*** [16]. **Pramanik *et al.*** have proposed a texture-based feature extraction approach called Block Variance (BV) that depends on asymmetric temperature distribution. Additionally, they have used a set of statistical features including the skewness, kurtosis, and the mean. Also, they have employed a feed-forward neural network with gradient decent training rule to perform the classification system. Their system has been evaluated using thermal breast image dataset that consists of sixty benign thermal breast images and forty malignant thermal breast images. **Sathish *et al.*** [15] have proposed a fully automated segmentation method for segmenting right and left breast for breast cancer detection using thermograms. The proposed segmentation method employs the breast's shape features besides to the polynomial curve fitting. In order to differentiate between normal and abnormal breast, histogram and GLCM based features are extracted. The classification step is performed using SVM-RBF classifier. The proposed system is evaluated using a thermal breast image dataset that consists of fourty normal thermograms and fourty abnormal thermograms. Additionally, **Hossam *et al.*** [16] have proposed a system for normal and abnormal thermal breast images. The proposed system starts with segmenting the Region of Interest automatically. Then, the boundaries of the segmented areas are enhanced using the Hough transform algorithm in addition to another approach for enhancing the contrast of the detected boundaries. Then, GLCM based features are extracted from the segmented images. Finally, the classification step is performed using two classifiers namely, SVM classifier and multi-layer perceptron with back-propagation classifier. The proposed system has been evaluated using a thermal breast images dataset that consists of 90 normal breast images and 110 abnormal breast images. The SVM classifier has achieved a better results compared to the other classifier. The performance metrics of the different methods are shown in Table 2.

From the obtained results, it is observed that using all of the extracted features leads to great accuracy where a number of important features can be eliminated during the feature selection step. However, regarding the recognition time, we can notice that PSO-MSVM showed exactly the same accuracy compared to the original system with less recognition time and ACO-MSVM take less recognition time with less accuracy which is reasonable where the size of the feature vectors can greatly the time needed for the classification process. In other words, the larger the size of the feature vector is, the larger the time needed to perform the classification process.

On the other hand, despite working on a multi-class classification problem rather than a binary classification problem and using a larger dataset compared with **Pramanik *et al.*** [14] and **Sathish *et al.* [15]**, it is noticed that all the proposed methods provides better performance than their works regarding the different performance metrics. The superiority of the proposed work over these works occurs due to using two types of features namely, statistical features (GLCM based features) and texture features (bag-of-words features) rather than using a single type of features. Also, the proposed original system and PSO-MSVM are better than the work proposed by **Hossam *et al.*** [16], which classify thermal images into only two classes, in terms of classification accuracy. Once again, the superiority of proposed system is due to using two types of features instead of using one type.

6. Conclusion

Breast cancer is a main cause of death among women. However, the early discovery of this disease can incredibly enhance the survival rates. In this paper, a thermography-based breast cancer detection method has been proposed. The proposed method has been mainly designed to distinguish between three different classes of thermography breast image: normal, benign, and malignant. The proposed method includes four main steps: preprocessing, feature extraction, feature selection, and classification. In the feature extraction step, two

groups of features including bag of features and GLCM based features are extracted from the thermography images. In the feature selection step, two swarm intelligence-based techniques namely ACO and PSO have been used to select the most discriminating features and exclude the rest of features to reduce the dimensionality of the feature vectors. Finally, in the classification step, a modified version of SVM called Multi-class SVM with polynomial kernel function has been used for building the classifier. The proposed methods have been evaluated using a standard thermography breast image dataset. The obtained results have shown the superiority of original method (without feature selection step) with 97.14% accuracy, 98% sensitivity, and 98.6% specificity, followed by PSO-MSVM and ACO-MSVM, in order. On other side, PSO-MSVM and ACO-MSVM required a shorter time in order to perform the recognition process compared to the original method. The recognition time is shorter in PSO-MSVM and ACO-MSVM because the length of feature vectors in these approaches is smaller compared to the length of the feature vectors in the original method due to the feature selection step. In addition, the obtained results have shown the superiority of the proposed approaches over other recently proposed approaches despite working on three classes rather than two classes and with a larger dataset. In the future, the proposed work can be extended by proposing an efficient automated segmentation approach to extract the ROI from thermal breast images automatically. In addition, deep learning approaches can be employed for feature extraction and/or classification steps.

References

[1] G. Sayed, M. Soliman, and A. Hassanien, "Bio-inspired Swarm Techniques for Thermogram Breast Cancer Detection", *In Medical Imaging in Clinical Applications*, *Springer International Publishing*, Vol. 13, No. 1, pp. 487-506, 2016.

[2] A. Bhattacherjee, S. Roy, S. Paul, P. Roy, N. Kausar, and N. Dey, "Classification Approach for Breast Cancer Detection Using Back Propagation Neural Network: A Study", *Biomedical Image Analysis and Mining Techniques for Improved Health Outcomes*, *IGI Global*, pp. 210-221, 2015

[3] A. Mert, N. Kılıç, E. Bilgili, and A.Akan, "Breast cancer detection with reduced feature set", *Computational and mathematical methods in medicine*, Vol. 2015, 2015.

[4] B. Krawczyk and G. Schaefer, "Effective Imbalanced Classification of Breast Thermogram Features", In: *Proc. of Springer International Conf. on Pattern Recognition and Machine Intelligence*, pp. 535-544, 2015.

[5] T. Gaber, G. Ismail, A. M. Anter, M. Soliman, M. Ali, N. Semary, A. Hassanien, and V. Snasel, "Thermogram Breast Cancer Prediction Approach Based on Neutrosophic Sets and Fuzzy C-means Algorithm", In: *Proc. of the IEEE Conf. on Engineering in Medicine and Biology Society*, pp. 4254-4257, 2015.

[6] A. E. Lashkari, F. Pak, and M. Firouzmand, "Full Intelligent Cancer Classification of Thermal Breast Images to Assist Physician in Clinical Diagnostic Applications", *Journal of Medical Signals and Sensors*, Vol. 6, No. 1, pp.12-24, 2016.

[7] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and Malignant Breast Tumors Classification Based on Region Growing and CNN Segmentation", *Expert Systems with Applications*, Vol. 42, No. 3, pp. 990-1002. 2015.

[8] K. Vaidehi and T. S. Subashini, "Automatic Characterization of Benign and Malignant Masses in Mammography", In: *Proc. of the International Conf. on Information and Communication Technologies*, pp.1762-1769, 2015.

[9] M. Taheri, G. Hamer, S. H. Son, and S. Y. Shin, October. Enhanced breast cancer classification with automatic thresholding using SVM and Harris corner detection", In: *Proc. of the ACM International Conf. on Research in Adaptive and Convergent Systems*, pp. 56-60, 2016.

[10] S. Thawkar and R. Ingolikar, "Automatic Detection and Classification of Masses in Digital Mammograms", *International Journal of Intelligent Engineering and Systems*, Vol.10, No.1, pp. 65-74, 2017.

[11] B. Krawczyk and G. Schaefer, "Breast Thermogram Analysis Using Classifier Ensembles and Image Symmetry Features", *IEEE Systems Journal*, Vol. 8, No. 3, pp. 921-928, 2014.

[12] W. Oleszkiewicz, P. Cichosz, D. Jagodzinski, M. Matysiewicz, L. Neumann, R. M. Nowak, and R. Okuniewski, "Application of SVM Classifier in Thermographic Image Classification for Early Detection of Breast Cancer", In: *Proc. of the IEEE-SPIE Joint Symposium on Photonics, Web Engineering, Electronics for Astronomy and High Energy Physics Experiments*, 2016.

[13] G. Schaefer, T. Nakashima, M. Zavisek, Y. Yokota, A. Drastich, and H. Ishibuchi, "Breast Cancer Classification Using Statistical Features and Fuzzy Classification of Thermograms", In: *Proc. of the IEEE Conf. on Fuzzy Systems*, pp. 1-5, 2007.

[14] S. Pramanik, D. Bhattacharjee, and M. Nasipuri, "Texture Analysis of Breast Thermogram for Differentiation of Malignant and Benign Breast", In: *Proc. of the IEEE Conf. on Advances in Computing, Communications and Informatics*, pp. 8-14, 2016.

[15] D. Sathish, S. Kamath, K. Prasad, R. Kadavigere, and R. J. Martis, "Asymmetry Analysis of Breast Thermograms Using Automated Segmentation and Texture Features", *Signal, Image and Video Processing*, Vol. 11, No. 4, pp 745-752, 2017.

[16] A. Hossam, H. M. Harb, and H. M. Abd El Kader, "Automatic Image Segmentation Method for Breast Cancer Analysis Using Thermography", *Journal of Engineering Sciences*, Vol. 46, No. 1, pp. 12-32, 2018.

[17] T. Botterill, S. Mills, and R. Green, "Speeded-up Bag-of-words Algorithm for Robot Localisation Through Scene Recognition", In: *Proc. of the IEEE International Conf. on Image and Vision Computing*, pp. 1-6, 2008.

[18] A. Alfanindya, N. Hashim, and C. Eswaran, "Content Based Image Retrieval and Classification Using Speeded-up Robust Features (SURF) and Grouped Bag-of-visual-words (GBoVW)", In: *Proc. of the IEEE International Conf. on Technology, Informatics, Management, Engineering, and Environment*, pp. 77-82, 2013.

[19] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural Features for Image Classification", *IEEE Transaction on Systems , Man and Cybernetics*, Vol. 3, No. 6, pp. 612–621, 1973.

[20] R. Rahimeh, M. Jafari, S. Kasaei, and P. Keshavarzian, "Benign and Malignant Breast Tumors Classification Based on Region Growing and CNN Segmentation", *Expert Systems with Applications*, Vol. 42, No. 3, ppt. 990-1002, 2014.

[21] M. A. Tawfeek, A. El-Sisi, A. E. Keshk, and F. A. Torkey, "Cloud task scheduling based on ant colony optimization", In: *Proc. of the IEEE International Conf. on Computer Engineering & Systems*, pp. 64-69 , 2013.

[22] O. Fallahzadeh, Z. Dehghani-Bidgoli, and M. Assarian, "Raman Spectral Feature Selection Using Ant Colony Optimization for Breast Cancer Diagnosis", *Lasers in Medical Science*, Vol. 33, No. 8, pp.1799-1806, 2018.

[23] B. Xue, M. Zhang, and W. N. Browne, "Particle Swarm Optimization for Feature Selection in Classification: A Multi-objective Approach", *IEEE Transactions on Cybernetics*, Vol. 43, No. 6, pp. 1656-1671, 2013.

[24] J. Kennedy and R. C. Eberhart, "A Discrete Binary Version of The Particle Swarm Algorithm", In: *Proc. of the IEEE International Conf. on Systems, Man, and Cybernetics, Computational Cybernetics and Simulation*, Vol. 5, pp. 4104-4108, 1997.

[25] http://www.mathworks.com/matlabcentral/fileexchange/33170-multi-class-support-vector-machine, Last Access: 18-8-2017.

[26] L. F. Silva, D. C. M. Saade, G. O. Sequeiros, A. C. Silva, A. Paiva, R. S. Bravo, and A. Conci, "A New Database for Breast Research with Infrared Image", *Journal of Medical Imaging and Health Informatics*, Vol. 4, No. 1, pp. 92–100, 2014.